

## RESEARCH ARTICLE

# Composition and origin of lung fluid proteome in premature infants and relationship to respiratory outcome

Philip L. Ballard<sup>1\*</sup>, Juan Osés-Prieto<sup>2</sup>, Cheryl Chapin<sup>1</sup>, Mark R. Segal<sup>3</sup>, Roberta A. Ballard<sup>1</sup>, Alma L. Burlingame<sup>2</sup>

**1** Department of Pediatrics, University of California, San Francisco, San Francisco, California, United States of America, **2** Department of Chemistry and Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, California, United States of America, **3** Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, California, United States of America

\* [Phil.ballard@ucsf.edu](mailto:Phil.ballard@ucsf.edu)



## OPEN ACCESS

**Citation:** Ballard PL, Osés-Prieto J, Chapin C, Segal MR, Ballard RA, Burlingame AL (2020) Composition and origin of lung fluid proteome in premature infants and relationship to respiratory outcome. PLoS ONE 15(12): e0243168. <https://doi.org/10.1371/journal.pone.0243168>

**Editor:** Olivier Baud, Hopital Robert Debre, FRANCE

**Received:** June 10, 2020

**Accepted:** November 13, 2020

**Published:** December 10, 2020

**Copyright:** © 2020 Ballard et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** The mass spectrometry proteomics files have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository (<http://www.ebi.ac.uk/pride>) with the dataset identifier PXD021966.

**Funding:** This study was funded by grants from the NIH National Heart Lung Blood Institute R21HL129910 (PLB) and the Department of Pediatrics University of California, San Francisco (PLB). ALB is director of the Bio-Organic Biomedical Mass Spectrometry Resource at UCSF, which is supported by the National Institute of

## Abstract

### Background

Infants born at extremely low gestational age are at high risk for bronchopulmonary dysplasia and continuing lung disease. There are no early clinical biomarkers for pulmonary outcome and limited therapeutic interventions.

### Objectives

We performed global proteomics of premature infant tracheal aspirate (TA) and plasma to determine the composition and source of lung fluid proteins and to identify potential biomarkers of respiratory outcome.

### Methods

TA samples were collected from intubated infants in the TOLSURF cohort before and after nitric oxide treatment, and plasma was collected from NO CLD infants. Protein abundance was assayed by HPLC/tandem mass spectrometry and Protein Prospector software. mRNA abundance in mid-gestation fetal lung was assessed by RNA sequencing. Pulmonary morbidity was defined as a need for ventilatory support at term and during the first year.

### Results

Abundant TA proteins included albumin, hemoglobin, and actin-related proteins. 96 of 137 detected plasma proteins were present in TA ( $r = 0.69$ ,  $p < 0.00001$ ). Based on lung RNAseq data, ~88% of detected TA proteins in injured infant lung are derived at least in part from lung epithelium with overrepresentation in categories of cell membrane/secretion and stress/inflammation. Comparing 37 infants at study enrollment (7–14 days) who did or did not develop persistent pulmonary morbidity, candidate biomarkers of both lung (eg., annexin A5) and plasma (eg., vitamin D-binding protein) origin were identified. Notably, levels of free hemoglobin were 2.9-fold ( $p = 0.03$ ) higher in infants with pulmonary morbidity. In time

General Medical Sciences (P41 GM103481). Ikaria Inc. (now part of Mallinckrodt Pharmaceuticals <https://www.mallinckrodt.ca/>)Biotech) and ONY Inc. (<https://www.onybiotech.com/>) provided drug for the conduct of the parent TOLSURF trial, but neither company had input into study design, data analysis, data interpretation, decision to publish or manuscript preparation. No authors received salary support from either commercial company.

**Competing interests:** This study was funded by grants from the NIH National Heart Lung Blood Institute R21HL129910 (PLB) and the Department of Pediatrics University of California, San Francisco (PLB). ALB is director of the Bio-Organic Biomedical Mass Spectrometry Resource at UCSF, which is supported by the National Institute of General Medical Sciences (P41 GM103481). Ikaria Inc. (now part of Mallinckrodt Pharmaceuticals <https://www.mallinckrodt.ca/>)Biotech) and ONY Inc. (<https://www.onybiotech.com/>) provided drug for the conduct of the parent TOLSURF trial, but neither company had input into study design, data analysis, data interpretation, decision to publish or manuscript preparation. No authors received salary support from either commercial company and there are no relevant declarations for either commercial funder relating to employment, consultancy, patents, products in development, or marketed products. This does not alter our adherence to PLOS ONE policies on sharing data and materials.

course studies, hemoglobin decreased markedly in most infants after enrollment coincident with initiation of inhaled nitric oxide treatment.

## Conclusions

We conclude that both lung epithelium and plasma contribute to the lung fluid proteome in premature infants with lung injury. Early postnatal elevation of free hemoglobin and heme, which are both pro-oxidants, may contribute to persistent lung disease by depleting nitric oxide and increasing oxidative/nitrative stress.

## Introduction

Infants born prematurely are at risk for respiratory failure and for bronchopulmonary dysplasia (BPD), a disorder that can continue into childhood and is characterized by hypoxemia, pulmonary hypertension, wheezing and/or later asthma, requiring chronic respiratory medications and hospitalizations. The pathogenesis of BPD and associated pulmonary sequelae is not fully defined but includes lung structural and biochemical immaturity with surfactant deficiency and dysfunction, low antioxidant and immune defenses, plus exposure to hypoxia/hyperoxia and barotrauma from ventilator support; these insults damage lung epithelium, promote pulmonary edema, and elicit sterile inflammation that can be compounded by infection [1, 2]. Currently, therapeutic options for the prevention and treatment of BPD are limited and have not substantially affected the incidence of disease that remains high for the very premature infants [3].

Infants with more severe lung disease are often intubated in order to provide mechanical ventilation and airway pressure, which provides access to tracheal aspirates (TA) after saline instillation. Previous studies have analyzed constituents of lung fluid such as cytokines and surfactant components with the goal of identifying potential biomarkers for later respiratory outcome and understanding the pathogenesis of infant lung disease [4, 5]. While a number of associations between lung fluid constituents and outcome have been reported, none are in use clinically to assess risk, guide interventions or serve as a proxy for long-term outcome in clinical trials.

The continuing development of omic technologies to profile gene and protein expression and metabolite levels provides additional opportunities to explore the composition of lung fluid related to development and disease. Numerous studies have reported disease-related changes in the proteome of bronchoalveolar fluid (BALF) in adults [6–12], but information for lung fluid of infants is limited to one study [13]. In the current study, we performed a global proteomic analysis on TA samples collected from premature infants enrolled in the Trial of Late Surfactant (TOLSURF) [14, 15]. The objectives of our study were to describe the proteome of infant lung fluid, determine the origins of lung fluid proteins, and identify proteins as potential biomarkers of later respiratory status. We report quantitative data for protein levels in lung fluid and plasma, and we investigate the origin of proteins using lung transcriptomic data and proteomic data from human fetal lung and infant plasma. Comparing infants with and without lung disease persisting through the first year, we identified candidate proteins that have biological plausibility for contributing to lung injury and may serve as early biomarkers of later respiratory status.

## Materials and methods

### Study population

The Trial of Late Surfactant (TOLSURF, ClinicalTrials.gov, NCT01022580) was a blinded, randomized, sham-controlled trial performed at 25 US centers and designed to assess effects of

late surfactant treatments on respiratory outcome. The trial design, infant characteristics and effects of late surfactant treatment have been reported [14]. A total of 511 extremely low gestational age ( $\leq 28$  weeks) infants who required mechanical ventilation at 7–14 days were enrolled in a randomized, masked controlled trial at 25 U.S. centers. All infants received inhaled nitric oxide (iNO) beginning at enrollment, as described for the NO CLD trial (ClinicalTrials.gov, [NCT00000548](https://clinicaltrials.gov/ct2/show/study/NCT00000548)) [16], and either surfactant (calfactant/Infasurf) or sham installation every 1–3 days for a maximum of 5 doses while intubated. The primary outcome was survival at 36 weeks postmenstrual age without BPD as evaluated by physiological oxygen/flow reduction. 252 treated infants received surfactant and 259 controls received a sham instillation procedure. Clinical respiratory parameters were collected and respiratory severity score (RSS,  $\text{FiO}_2 \times$  mean airway pressure) was calculated. Survival without BPD did not differ between the treated and control groups at either 36 or 40 wk postmenstrual age, and there were no between-group differences in serious adverse events, comorbidities or severity of lung disease to 36 weeks. Respiratory outcomes at 1-year corrected age were assessed by quarterly questionnaires addressing use of respiratory medication, hospitalization for lung disease, and home respiratory support as reported [15].

For the current study, severe respiratory outcome was defined as a diagnosis of BPD at 36 wk and 40 wk postmenstrual age plus post-discharge respiratory support reported in 3 of 4 quarters and is designated as pulmonary morbidity (PM). This definition combines 40-wk BPD and persistent pulmonary morbidity during the first year as previously defined by Keller et al. for TOLSURF infants [15]. No/mild lung disease was defined as no BPD at 40 wk and no first-year pulmonary morbidity (No PM). Thus, the PM/No PM definition describes infants at extremes of respiratory disease. Some of the other infants who were studied over time for TA protein abundance had an intermediate respiratory outcome: i.e., BPD at 40 weeks and  $\leq 2$  quarters with respiratory support or no BPD at 40 wk and post-discharge respiratory support reported in  $\geq 1$  quarter.

Biorepository plasma samples from 8 infants of the NO CLD Trial were used for proteomic assay as described for TA samples. Infants in NO CLD were similar to the TOLSURF cohort with regard to gestational age ( $25.3 \pm 1.3$  wk) and requirement for intubation and ventilatory support at enrollment, however most were older at blood collection ( $41 \pm 16$  d) than for TA collection in TOLSURF infants ( $9 \pm 2$  d).

Enrolled infants from both TOLSURF and NO CLD had written, informed consent by a parent and all studies were approved by all Institutional Review Boards at the following institutions: University of Arkansas—Arkansas Children’s Hospital Little Rock, AR; Alta Bates Medical Center, Berkeley, CA; Oakland Children’s Hospital, Oakland, CA; University of California, San Francisco, San Francisco, CA; Wolfson Children’s Hospital and Shands Hospital, Jacksonville, FL; Florida Hospital for Children, Orlando, F; All Children’s Hospital, St Petersburg, FL; Northwestern Memorial Hospital, Chicago, IL; Children’s Memorial Hospital, Chicago, IL; Children’s Hospital and Clinics of Minnesota- Minneapolis, Minneapolis, MN; University of Minnesota Medical School, Minneapolis, MN; Children’s Hospital and Clinics of Minnesota—St Paul, St Paul, MN; Children’s Mercy Hospital, Kansas City, MO; Women’s and Children’s Hospital of Buffalo, Buffalo, NY; Stony Brook University Medical Center, Stony Brook, NY; Wake Forest University- Forsyth Hospital and Brenner Hospital, Winston-Salem, NC; Medical University of South Carolina, Charleston, SC; University of Tennessee Memphis- Memphis Medical Center, Memphis, TN; University of Texas Houston Health Science Center, Houston, TX; Texas Children’s Hospital, Houston, TX; University of Washington, Seattle, WA; Children’s Hospital of Philadelphia, Philadelphia PA; University of Pennsylvania, Philadelphia PA; Case Western Reserve University, Cleveland OH; Rainbow Babies and Children’s Hospital, Cleveland OH; Brigham and Women’s Hospital, Boston MA; Children’s Hospital, Boston

MA; Cedars-Siani Medical Center, Los Angeles CA; Columbus Children's Hospital, Columbus OH; Schneider Children's Hospital, New Hyde Park NY; Westchester Medical Center, Valhalla NY; University of Utah Hospital and Clinics, Salt Lake City, UT; Primary Children's Medical Center, Salt Lake City, UT.

### Tracheal aspirate samples

In the TOLSURF trial, specimens of TA were collected from infants at enrollment (postnatal age 7–14 d) before treatment with surfactant or sham instillation and before starting iNO; additional samples were collected from infants before each dose of surfactant/sham. To characterize the TA proteome and investigate potential biomarkers for respiratory outcome, we selected TA samples at enrollment for 37 infants who had either PM ( $n = 20$ ) or No PM ( $n = 17$ ), as defined above, and were matched by study site to reduce variability in clinical management practices that might influence outcome. We used sequential TA samples from an additional 24 infants to investigate changes in protein abundance over time after enrollment. The total group of 61 infants was similar to the total TOLSURF cohort with gestational age of  $25.7 \pm 1.3$  wk, birth weight  $734 \pm 183$  g, 51% male, 43% white, 32% black, and 19% Hispanic.

The TA sampling procedure involved 2 instillations of saline into the trachea, brief ventilation, and suction to recover saline containing lung epithelial lining fluid. Samples were centrifuged at 500 $\times$ g for 5 min to remove cells and the supernatant stored at  $-70$  C in the presence of protease inhibitors.

### mRNA analysis

RNA was isolated from mid-gestation (19–23 wk gestation) specimens of human fetal lung that were collected with institutional review board approval from elective pregnancy terminations in healthy consenting women. mRNA sequencing and analysis were performed by the UCSF SABRE Functional Genomics Core using Hi-Seq 2500 machines (Illumina, San Diego, CA) as described [17].

### Fetal lung explant culture

To examine proteins in the medium of cultured lung tissue, explants of human fetal lung were prepared and cultured in serum-free Waymouth's medium as described [18]. Fresh medium without additives was added after 24 h of culture and then collected 48 h later for proteomic analysis as described for TA samples.

### Protein analyses and statistics

Total protein in TA samples and culture medium was measured in duplicate by the Bradford assay (BioRad Laboratories, Hercules CA). Cell free hemoglobin was assayed in duplicate using Hemoglobin Human ELISA Kit (ab157707, Abcam, Burlingame, CA). For global proteomic analysis, aliquots of TA, culture medium or plasma were digested into peptides with trypsin and analyzed on a QExactive Plus mass spectrometer (Thermo Scientific, Waltham, MA) connected to a NanoAcquity™ Ultra Performance UPLC system (Waters, Milford MA). Peak lists were generated using PAVA in-house software [19] and searched against the human subset of the SwissProt database, using Protein Prospector [20]. The number of spectra identified as matches to peptides of a given protein were compiled, and these values (Peptide Spectral Matches, PSMs) were used for label-free quantitation of protein abundance in the samples. A relative abundance index (AI) value was calculated for each protein, which is the PSM spectral count divided by the size of the protein, divided by the total spectral counts in the sample. The

mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository (<http://www.ebi.ac.uk/pride>) with the dataset identifier PXD021966 [21].

For proteomic analysis, all the TA (or plasma samples) from each infant were run in the same batch to eliminate inter-assay variability. AI data are expressed as mean $\pm$ sd for each group. Relationships between levels of proteins in TA vs plasma or culture medium were evaluated by linear regression. For analysis of protein AI in the first TA sample from infants with and without PM, we performed log transformation and Student's t test (unadjusted p values); we also used the moderated or penalized version developed for large scale hypothesis testing situations as implemented in the R package limma, which has been demonstrated to be robust with respect to a variety of departures from normality (q values). We used  $p = 0.10$  as a cut-off value for these exploratory studies of differential abundance. For the derived value of total hemoglobin (sum of all subunits), we used Mann-Whitney analysis. Statistical evaluation of Gene Ontology enrichment process was performed through the Gene Ontology Consortium website [22].

Initial experiments examined the intra- and inter-assay variability for each identified TA protein using 2 TA samples and 4 replicates of each sample (Supplemental S1A and S1B Fig, respectively). The coefficients of variation were dependent on the peptide count of a protein; for proteins with a peptide count  $>3$ , the coefficients of variation were 18.5% and 32.2% for intra- and inter-assay variability. Based on these data we limited quantitative comparisons of protein abundance to proteins with a mean peptide count of  $>3$ .

## Results

### TA Proteins detected by MS:MS

Studies of TA samples from 37 infants at enrollment into TOLSURF detected a total of 1839 different proteins as defined by the presence of at least 1 unique peptide in at least 1 TA sample (S1 Table). 809 of these proteins were present in  $\geq 25\%$  of samples, and we considered this level of occurrence in infant TA as sufficient to qualify as a TA protein. Approximately half (408) of these proteins had a mean peptide count  $\geq 3$ , which was judged sufficient for quantitation of protein AI.

The 20 TA proteins with highest AI values are shown in Table 1. Albumin was the most abundant protein. Other proteins include 5 hemoglobin subunits, iron-transport/trafficking proteins *TF* and *LCN2*, 3 immune-related proteins Ig chains and *C3*, 4 actin (*ACTB*, *ACTG1*) and actin-related (*TMSB4X*, *PFN1*) proteins, and lung epithelium secreted proteins (*SERPINA1*, *SFTPB*, *SCGB1A1*).

To explore which of the proteins detected in TA likely originated, at least in part, from the lung, we compared TA proteins with results by RNAseq for gene expression in late second trimester human fetal lung [17]; overall, mRNA was detected for 83% of detected TA proteins (S2 Table). By linear regression analysis of lung mRNA abundance versus TA protein AI, there was a significant slope but low positive correlation ( $r = 0.09$ ,  $p < 0.02$ ). Thus, of the proteins in TA detected by global proteomic analysis, the great majority are produced in lung cells and may represent, at least in part, proteins secreted into lung fluid and/or released secondary to lung injury.

In additional experiments using explants of fetal lung cultured in serum-free medium, we examined whether proteins corresponding to expressed mRNAs were released from cells [17]. Restricting the proteomic analysis to proteins with detectable lung mRNA, there were 315 proteins reliably detected and quantitated (S5 Table) in the culture medium. Protein abundance in the medium was significantly correlated with abundance in TA (Fig 1). Some apparent

Table 1. TA proteins of highest abundance.

Protein	Gene Symbol	Peptides Detected/Sample (mean±sd)	Protein AI x 10 <sup>8</sup> (mean±sd)	mRNA cpm (mean±sd)
Serum albumin	<i>ALB</i>	1110±451	178.5±47	ND
Hemoglobin subunit alpha	<i>HBA1</i>	83±141	60.1±87	10±3
Hemoglobin subunit gamma-1	<i>HBG1</i>	62±96	44.0±59	28±22
Hemoglobin subunit gamma-2	<i>HBG2</i>	61±103	43.5±50	219±30
Hemoglobin subunit beta	<i>HBB</i>	50±63	37.2±46	53±4
Thymosin beta-4	<i>TMSB4X</i>	14±14	37.1±55	1525±91
Ig kappa chain C region	<i>IGKC</i>	31±9	31.6±11	ND
Actin, cytoplasmic 1	<i>ACTB</i>	114±39	31.2±8	1924±231
Actin, cytoplasmic 2	<i>ACTG1</i>	112±39	30.8±8	1942±278
Uteroglobin	<i>SCGB1A1</i>	26±18	28.7±18	ND
Serotransferrin	<i>TF</i>	174±61	25.3±6	ND
Alpha-1-antitrypsin	<i>SERPINA1</i>	101±45	24.1±8	1625±1132
Histone H4	<i>HIST1H4A</i>	23±10	22.6±8	ND
Hemoglobin subunit delta	<i>HBD</i>	27±34	20.6±26	ND
Ig gamma-1 chain C region	<i>IGHG1</i>	62±19	20.1±7	ND
Surfactant-associated protein B	<i>SFTPB</i>	76±36	20.1±7	1716±1229
Profilin-1	<i>PFN1</i>	24±7	19.1±9	331±28
Ig lambda-2 chain C regions	<i>IGLC2</i>	18±8	18.6±9	ND
Complement C3	<i>C3</i>	296±98	17.9±5	3043±833
Neutrophil gelatinase-associated lipocalin	<i>LCN2</i>	32±16	17.5±11	234±188

Protein data from TA samples of 37 infants at 7–14 days, before receiving iNO and surfactant or sham, and mRNA data from mid-trimester human fetal lung. Clinical characteristics of the infants are presented in Table 3. cpm, counts per million.

<https://doi.org/10.1371/journal.pone.0243168.t001>

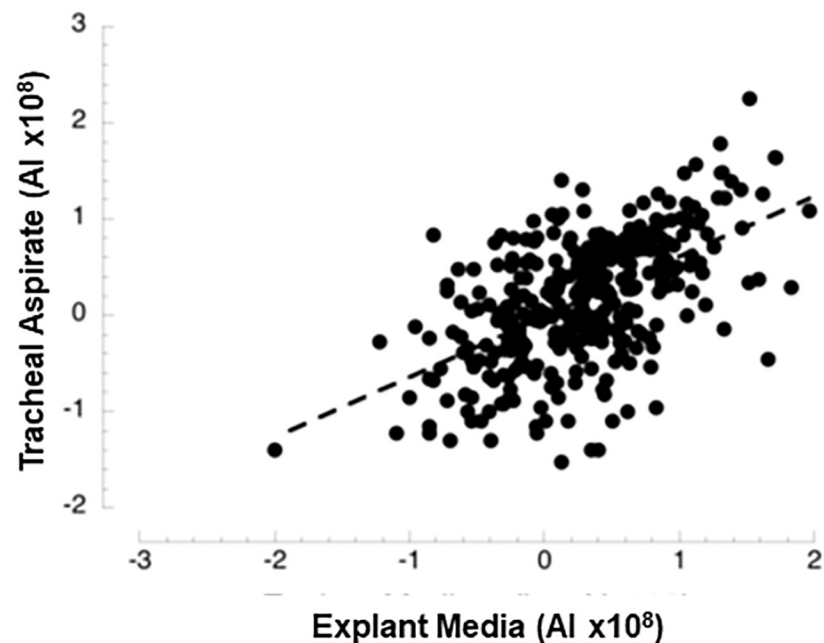


Fig 1. Comparison of abundance levels for proteins of TA and culture medium of human fetal lung explants. By linear regression analysis of log transformed data there is a highly significant positive correlation ( $r = 0.56$ ,  $p < 10^{-30}$ ).

<https://doi.org/10.1371/journal.pone.0243168.g001>

outlier proteins were observed including 4 hemoglobin subunits with relatively high TA abundance and 3 proteins with low TA abundance relative to the medium: heme binding protein 2, tissue inhibitor metalloprotease 1 and a histone.

Analysis of TA proteins by Gene Ontology revealed highly significant overrepresentations of specific groups of proteins by both cellular component and biological process categories with enrichment ranging between 2- and 9-fold as summarized in Table 2. Highest categories of localization were for proteins of focal adhesions in the extracellular matrix, membrane rafts, cytoskeleton, and secreted/extracellular organelles. By function, enriched categories included proteins primarily related to secretion, protein processing and inflammation with a 9-fold enrichment of proteins related to neutrophil activation and function. In general, these findings are consistent with both normal lung epithelial cell function (e.g., secretion of surfactant) and inflammatory plus cell damage responses to lung injury in the immature lung [23]. A full listing of gene ontology results is given in S3 and S4 Tables.

### Plasma proteomics

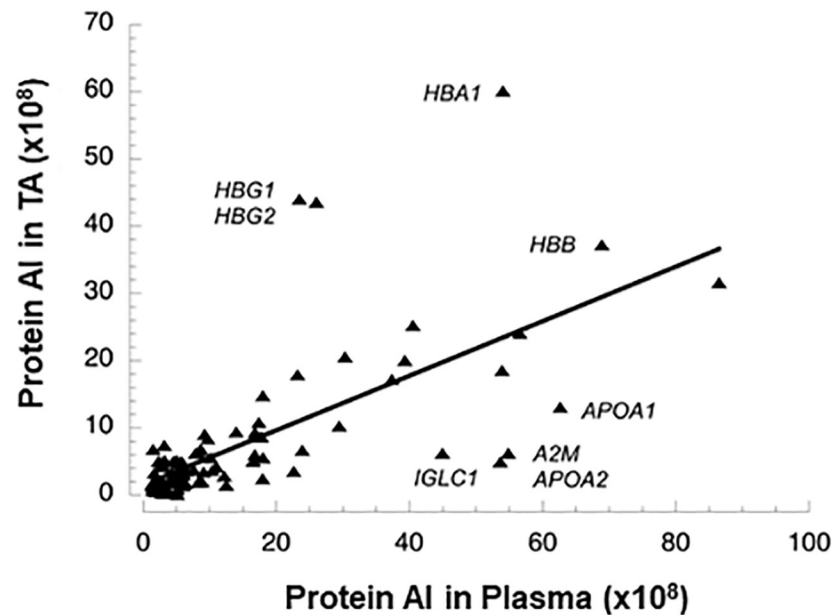
To directly investigate the contribution of circulating proteins to the infant TA proteome, we performed proteomic analysis of plasma using banked samples from intubated premature infants in the NO CLD trial (n = 8, mean gestational age 25.3 wk). A total of 435 proteins were detected in at least 25% of the plasma samples (S6 Table). Of these, 137 proteins were consistently present with a mean peptide count of  $\geq 3$ , providing a reliable value for AI, and of these, 96 (70%) were also consistently detected in TA samples. For these proteins, there was a close association between AI levels in TA and plasma (Fig 2,  $r = 0.69$ ,  $p < 0.00001$ ) excluding albumin, which was many-fold more abundant than other proteins. The median ratio of AI for all proteins in TA vs plasma was 0.52 with some outlier values: hemoglobin subunits alpha and

**Table 2. Cellular localization and biological process of TA proteins by gene ontology.**

GO Cellular Component	GO Proteins	TA Proteins	Expected Proteins	Fold Enrichment
Focal adhesion/cell-substrate junction	404	104	12	8.5
Extracellular exosome/vesicle/organelle	2109	488	64	7.6
Secretory granule/vesicle	911	183	28	6.7
Lysosome	670	94	20	4.6
Actin cytoskeleton	490	73	15	4.9
Membrane raft/microdomain	304	41	9	4.4
Intrinsic/integral component of membrane	5568	114	169	0.67
<b>GO Biological Process</b>				
Neutrophil function	496	141	15	9.4
Protein folding	182	31	6	8.2
Exocytosis/secretion	881	177	27	6.8
Inflammatory response	343	36	10	3.7
Response to cytokine	687	65	21	3.4
Response to stress	1349	105	41	3.3
Immune system/response/activation	1258	118	38	3.1
Protein transport/localization	966	71	29	2.9
Regulation of protein metabolic process	1014	79	31	2.9
Transcription-related	2564	37	78	0.47

GO categories for TA proteins with detectable mRNA in fetal lung. p values with Bonferonni correction ranged from 1.0E-79 to 1.0E-6 except for the 2 underrepresented categories (intrinsic/integral component of membrane and transcription-related) with  $p < 0.005$ .

<https://doi.org/10.1371/journal.pone.0243168.t002>



**Fig 2. Comparison of abundance levels for proteins present in both TA and plasma of premature infants.** By regression analysis  $r = 0.69$  ( $p < 0.00001$ ) excluding albumin as an outlier, which had AI values of 179 and 315 for TA and plasma, respectively. Gene symbols are shown for some outlier proteins.

<https://doi.org/10.1371/journal.pone.0243168.g002>

gamma 1 and 2 had higher TA:plasma ratios (2.2 to 4.5), and Ig lambda chain C, alpha2 macroglobulin and apolipoproteins A1 and A2 had TA:plasma ratios  $< 0.2$ . The TA abundance of the 11 outlier proteins was not related to protein molecular weight, which could affect trans-endothelial passage, nor was there a correlation between molecular weight and TA:plasma ratio for all proteins ( $r = 0.1$ ,  $p = 0.33$ ). The finding that 96 of 809 proteins detected in TA are also found in plasma suggests that influx of proteins into air spaces from the circulation can contribute up to  $\sim 12\%$  of detected lung fluid proteins. By relative abundance, 15 of the top 20 most abundant proteins of TA (Table 1) are plasma proteins.

### TA proteins and respiratory outcome

One objective of this study was to identify TA proteins that differed in abundance by two extremes of respiratory outcome. The 37 infants used for study of TA protein composition and origin were selected by matching infants by site. 20 PM infants and 17 No PM infants that had TA samples at enrollment were available for study; key demographics are shown in Table 3. All infants were at high risk for BPD based on TOLSURF entry criteria. Mean gestational ages were  $\sim 26$  weeks with mean birth weights of 685 and 748 g, respectively. The PM group had somewhat more males and Caucasians and a higher RSS score than the NO PM infants, but these differences were not significant; however, the number of blood transfusions was 2-fold higher ( $p < 0.01$ ) in PM infants.

Table 4 shows 20 TA proteins selected for apparent difference (unadjusted  $p < 0.10$ ) in abundance by respiratory outcome; 11 proteins were decreased (fold range 0.44–0.83) and 9 were increased (fold range 1.4–4.4) in PM vs No PM infants. A full list of protein abundance by outcome is shown in S7 Table. mRNA was not detected in fetal lung for 7 of the proteins, consistent with plasma as the source. Due to the limited number of infants studied and the large number of proteins assayed, none of the differences reached  $q < 0.10$  by false discovery



**Table 3. Characteristics of extreme outcome infants in study of TA proteins.**

	No PM	PM
N	17	20
Gestational age (wk)	25.9±1.2	25.5±1.4
Birth weight (g)	748±199	685±177
Male/female	6/11	13/7
Race (Caucasian/AA/Hispanic/Other)	5/6/5/1	11/6/2/1
Age enrollment/TA collection (days)	9.3±2.7	8.9±2.1
Number of clinical sites	12	13
Late surfactant therapy	9 (52.9%)	13 (65%)
RSS at TA collection	2.6±0.9	3.3±1.8
Blood transfusions (total number)	6.1±5.7	11.5±6.2 <sup>+</sup>
Patent ductus arteriosus—treated	5 (29.4%)	2 (10%)
Culture-positive sepsis	2 (11.8%)	0 (0%)
Severe pulmonary hemorrhage	1 (5.9%)	1 (5%)
Survive without BPD36wk (yes/no)	16/1	0/20
Survive without BPD40wk (yes/no)	17/0	0/20
Pulmonary morbidity first year (none/severe)*	17/0	0/20

Data for patent ductus arteriosus, sepsis and pulmonary hemorrhage are for the interval between birth and enrollment.

\*Severe is persistent pulmonary morbidity defined as respiratory support (medication, home respiratory support, hospitalization) in  $\geq 3$  of 4 quarters by questionnaire during the first year, and none is no reports of respiratory support during first year [15].

<sup>+</sup>  $p < 0.01$  by two-sided unpaired t test; AA, African American; RSS, respiratory severity score ( $\text{FiO}_2 \times \text{MAP}$ ).

<https://doi.org/10.1371/journal.pone.0243168.t003>

rate. Nevertheless, these are candidate proteins of interest for association with respiratory outcome. Of particular interest, all detected subunits of hemoglobin were higher in PM infants (2.6- to 4.4-fold), as was the derived value for total hemoglobin (2.9-fold,  $p = 0.029$ ), whereas levels were  $\sim 1.0$  for the major plasma proteins albumin and total immunoglobulins. For samples with higher levels of hemoglobin, there was a strong, positive correlation between levels determined by MS:MS versus immunoassay (Fig 3).

### TA time course

We analyzed protein abundance in additional, consecutive TA samples collected for a separate group of 24 infants with at least 3 longitudinal TA samples (7 PM, 9 No PM, 8 with intermediate outcome) between 7 and 21 days (3–5 TAs per infant over  $8 \pm 2$  days). For the 13 non-hemoglobin proteins that are candidates for differential abundance by respiratory outcome (Table 4), there was no consistent temporal change in AI levels with sequential samples over time. Similarly, serum albumin was present in every TA sample with a relatively narrow AI range (99–309) and no consistent temporal pattern. By contrast, the level of total hemoglobin decreased over time after trial enrollment in 17 of 21 infants with detectable hemoglobin to a mean level in the final TA sample of  $15 \pm 10\%$  of the first value; there was no consistent change in hemoglobin in 4 infants (3 PM infants and 1 intermediate outcome infant—last sample mean 99% of first sample). The decrease in total hemoglobin with time was similar for No PM and PM infants ( $14 \pm 15\%$  vs  $17 \pm 10\%$  of the first value, respectively). Fig 4 illustrates decreasing total hemoglobin compared to albumin, total IgG and total peptide count for a PM infant with an initial high level of total hemoglobin.

**Table 4. Selected proteins and relative abundance in TA of PM vs No PM infants.**

Protein	PM vs No PM		Fetal Lung Expression	
	Fold	AI	Protein	mRNA
Vitamin D-binding protein	0.77	10.7	1.2	<2
Gamma-enolase	0.71	2.1	2.3	192
Annexin A5	0.77	8.7	11.5	880
Complement C3	0.83	18.1	41.7	3043
Ubiquitin-like modifier-activating enzyme 1	0.62	0.5	<0.05	<2
Malate dehydrogenase, mitochondrial	0.48	0.6	2.6	113
Heterogeneous nuclear ribonucleoprotein A1	0.59	0.8	2.2	1179
Alpha-actinin-2	0.72	1.2	1.3	<2
Dipeptidyl peptidase 1	0.63	1.2	0.6	389
Tetranectin	0.63	1.8	2.2	<2
Purine nucleoside phosphorylase	0.52	1.5	<0.05	77
Vitamin K-dependent protein S	0.56	0.6	0.6	188
Rab GDP dissociation inhibitor beta	0.79	3.2	2.6	388
Alpha-1-acid glycoprotein 2	0.78	6.6	0.5	20
Beta-hexosaminidase subunit beta	0.58	0.5	0.4	118
Collagen alpha-1(VI) chain	0.53	0.6	1.5	145
Heat shock 70 kDa protein 6	1.4	1.6	1.2	61
Putative heat shock 70 kDa protein 7	1.55	1.7	1.8	9
Ras-related protein Rab-1B	1.96	1.1	<0.05	39
Histone H2A type 2-C	2.82	2.3	<0.05	<2
Hemoglobin subunit gamma-2	3.91	42.8	52	219
Hemoglobin subunit epsilon	2.77	5.1	<0.05	<2
Hemoglobin subunit gamma-1	2.85	43.3	51.4	28
Hemoglobin subunit beta	2.47	36.5	13.3	53
Hemoglobin subunit delta	2.36	20.2	<0.05	<2
Hemoglobin subunit zeta	3.45	2	1.9	<2
Hemoglobin subunit alpha	2.41	59.9	20.2	10
Total hemoglobin*	2.92	206.7	138.8	310

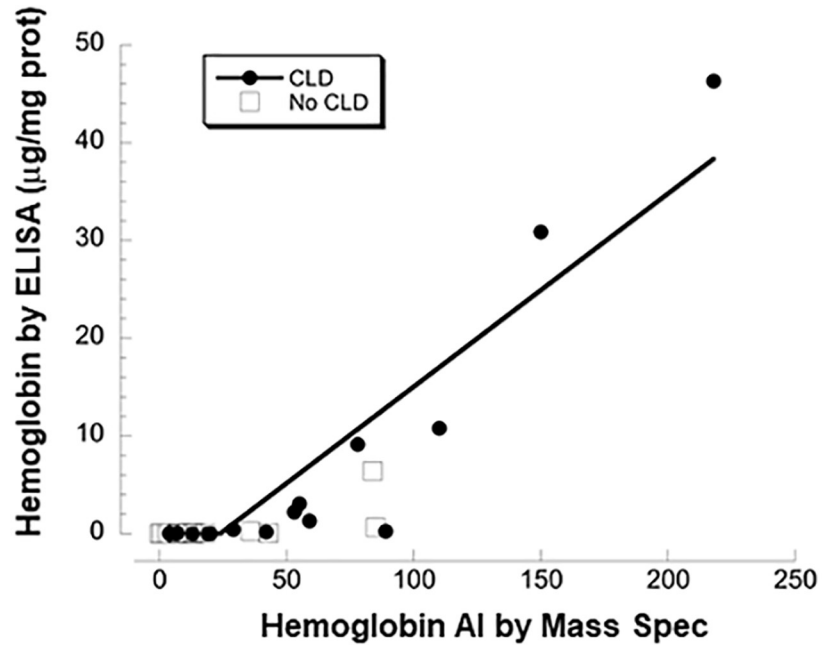
TA proteins with p value <0.10 (unpaired t test) comparing levels at enrollment of 37 infants with extreme respiratory outcomes. Levels of expression in lung are expressed as AI (protein abundance in explant medium) and mRNA as counts per million (cpm).

\*Calculated as sum of all subunits; p = 0.029 by Mann Whitney analysis.

<https://doi.org/10.1371/journal.pone.0243168.t004>

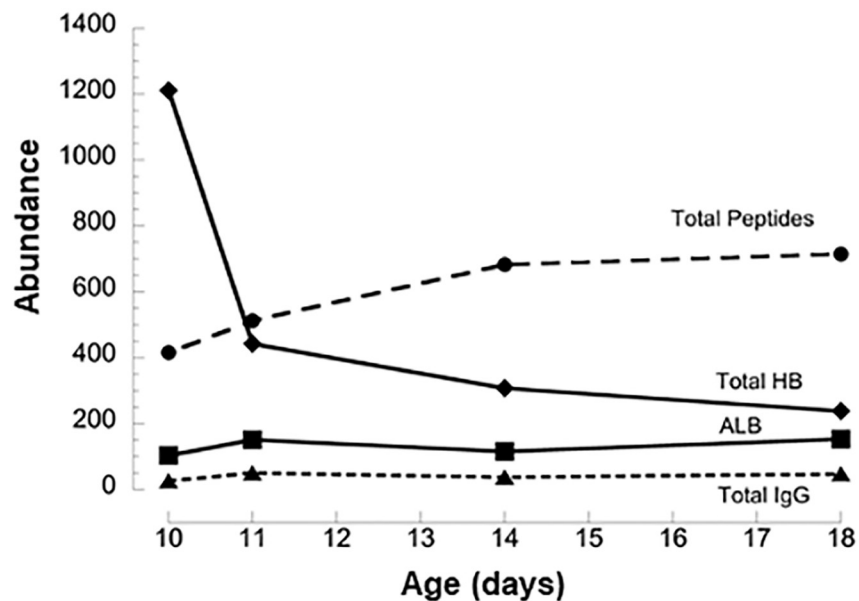
## Discussion

In this study, we used global proteomic analysis to provide new information related to TA proteins and their relative abundance in intubated premature infants. We also determined the proteome of premature infant plasma and examined both gene expression and presence of secreted proteins in human fetal lung. Assessing these complementary data, we report, to our knowledge, the first description of the origin of major proteins in TA of the premature infant. We conclude that up to 88% of the detected proteins in TA of these infants are likely derived primarily from lung epithelial cells via secretion and/or release secondary to injury to the extracellular matrix, cell membrane and cytoskeleton. Other TA proteins are primarily derived from plasma, with a strong correlation between relative abundance in plasma and TA. When the origin of TA protein is considered in terms of bulk amount, plasma contributes ~28% of total protein with serum albumin and free hemoglobin as major contributors. In comparison of the proteomes of two infant cohorts with extreme respiratory outcomes, candidate protein



**Fig 3. Comparison of abundance levels for total hemoglobin of TA by MS and ELISA assay.** Solid circles are infants with PM and open squares are No PM.  $n = 34$ ,  $r = 34$ ,  $p = 0.00001$  by linear regression. Except for lower values, results by immunoassay correlate with MS results.

<https://doi.org/10.1371/journal.pone.0243168.g003>



**Fig 4. Time course for TA levels of total peptides, total hemoglobin, albumin and total IgG.** Example of one infant with TA samples collected between 10 and 18 d postnatal age, showing decreases over time in hemoglobin, but not other major plasma proteins. Values are  $AI \times 10^3$  except for total peptide.

<https://doi.org/10.1371/journal.pone.0243168.g004>

biomarkers for PM of both lung and plasma origin were identified. The most notable of these was cell-free hemoglobin, which was elevated 2.9-fold in PM vs No PM infants at study enrollment.

Analysis of the infant TA proteome data by gene ontology revealed overrepresentation of proteins consistent with injury to the lung epithelium and inflammation. We did not have access to TA from control premature infants without lung disease, who are not intubated, but proteome data by a similar shotgun proteomic approach) are available for normal adult BALF [12]. Various families of proteins with relatively high abundance in infant TA were not detected in normal adult lung fluid: collagens, actins, heat shock proteins, histones, keratins, tubulins and subunits of fetal and embryonic hemoglobin. These differences between premature infant and adult lung could reflect a response to lung injury and/or developmental changes in expression levels.

Proteomic analyses of human lung fluid obtained by bronchopulmonary lavage has been reported for a variety of adult lung diseases including COPD [6], pulmonary sarcoidosis [7], idiopathic pneumonia syndrome [8], idiopathic pulmonary fibrosis [9], acute respiratory distress syndrome [10], ventilator-associated pneumonia [11] and HIV [12]. Most relevant to our study, in lung injury due to acute respiratory distress syndrome, Bhargava et al. [10] reported that non-survivors showed overrepresentation by gene ontology of proteins related to carbohydrate catabolism, disorganization of actin filaments and collagen biosynthesis. In a global proteomic study of adults with acute lung injury, Nguyen et al [12] found differential expression of nearly half of all detected proteins compared with normal controls.

To our knowledge, there has been only one previous report of proteomic analysis of infant tracheal aspirates. Magagnotti et al. [13] reported lower abundance of 2 calcium-binding proteins (*CAPS*, *CIB1*) and the chloride channel protein (*CLIC1*) for 6 of 12 infants who developed severe vs mild BPD. These proteins are expressed in fetal lung [17] and were detected in our proteomic study, however abundance in TA was not different by respiratory outcome.

The plasma proteome for adults and infants [24] has been previously described, but, to our knowledge, we provide the first comparison of infant plasma and lung fluid protein profile and evidence for the close proportional abundance in the two tissue compartments. This finding indicates that most plasma proteins, irrespective of their size, pass the endothelial/epithelial barrier and accumulate in lung fluid in the injured infant lung similar to the observation in adults with acute respiratory distress syndrome [25].

This study has limitations primarily related to methodology and the infant cohort. Proteomics by HPLC MS:MS provides a global survey but does not have the sensitivity of some other approaches such as immunoassay, which has documented the presence of numerous cytokines and chemokines in infant TA that we did not routinely detect [4]. Because plasma samples were not collected in TOLSURF, we used plasma from a separate cohort of infants (NO CLD) of similar demographics and severity of lung disease, however the plasma samples were collected later in the first month compared to the TA samples; future confirmation of relative abundance of proteins in TA versus plasma should involve matched TA:plasma samples from infants. Importantly, our findings are for infants with lung injury, and it is known from lamb studies that ventilator-induced lung injury results in a rapid and large change in protein composition of lung tissue, including extracellular proteins of epithelial origin [26]. Also, we compared TA proteins with proteins secreted from cultured fetal lung, which was not exposed to the insults associated with postnatal lung injury. In the biomarker component of our study we were limited in terms of the number of available infants with TA samples who had extreme respiratory outcomes. Because of this limitation we could not analyze our results by multivariate modeling to account for clinical predictors of lung disease, and thus our results for candidate proteins should be viewed as preliminary and will require validation in another, larger cohort.

The level of plasma free hemoglobin in healthy individuals is undetectable or low (<15 mg/dl). Levels can be many-fold higher with extra-corporeal membrane oxygenation and hemolytic anemias, but do not usually achieve the concentrations that we observed in plasma of infants with lung disease (~50% of albumin or ~1000 mg/dl). We also found high abundance of hemoglobin in TA, at levels similar those for albumin. By contrast, the level of hemoglobin alpha/beta was <5% of albumin in lung fluid of adults with acute lung injury [11, 12]. Thus, high levels of hemoglobin in lung fluid and plasma with lung injury may be unique to the premature infant, although further studies on this topic are needed.

The second observation regarding TA hemoglobin was the association of increased levels with later pulmonary morbidity. Of interest, Pereira-Fantini et al. [27] recently reported on changes in the plasma proteome in newborn premature lambs with exposure to 60 minutes of ventilator-induced lung injury. Similar to our results, three hemoglobins had the highest fold-increase (>2-fold) among 14 differentially abundant proteins of 125 total identified plasma proteins.

We have considered possible sources of increased cell-free hemoglobin in infant lung fluid: 1) intra-pulmonary bleeding and hemolysis; 2) increased synthesis/release of selected hemoglobin subunits (alpha, beta, gamma 1/2) from lung epithelial cells; 3) influx of cell-free plasma hemoglobin derived from blood transfusions; and 4) intrapulmonary hemolysis secondary to endothelium injury/dysfunction. Diagnosis of severe pulmonary hemorrhage was low (5–6%) and not different between outcome groups, however no clinical data are available for milder forms of hemorrhage. The second possibility (synthesis in lung) is supported by the finding of hemoglobin expression in human fetal lung (Table 4) and in lung type 2 cells of both fetal and adult human [23, 28] and by the report of increased hemoglobin gene expression and protein accumulation in rats exposed to hypoxia, which is a component of infant lung disease [29]. Cell-free hemoglobin and iron derived from transfused packed red cells, particularly with longer storage, may contribute to long-term lung disease based on the greater number of transfusions in PM vs No PM infants (Table 3) and on previous discussion [30]. The fourth possibility, hemolysis in the pulmonary circulation (and perhaps systemically) secondary to lung injury/oxidative stress, is supported by the observation of increased plasma hemoglobin subunits in lambs with ventilator-induced lung injury [27]. The finding of increased free hemoglobin with lung injury in two species supports further study of free hemoglobin (plasma and TA) as a potential early biomarker and monitoring approach for the severity of lung injury in infants.

Increased cell-free hemoglobin in lung fluid of infants, regardless of the sources, has biological plausibility for contributing to lung injury and affecting respiratory outcome. As an abundant protein, it could contribute to protein inhibition of surfactant activity, particularly when surfactant is deficient [5]. Second, as recently reviewed [31], cell-free hemoglobin is a pro-oxidant that would promote oxidative stress systemically and within airspaces through reactions with peroxides and nitric oxide (NO). Consumption of biologically available NO affects both vascular tone (eg., pulmonary arterial hypertension) and angiogenesis/repair processes after lung injury [32]. In studies with mice, tracheal instillation of cell free hemoglobin increased alveolar-capillary permeability, inflammatory cell influx and cytokine production more than albumin instillation [33]. Third, oxidation of hemoglobin via reaction with NO releases free heme that is toxic to membrane lipids and lipoproteins, inducing inflammation, cell injury and possibly surfactant dysfunction. Reduced amount of hemopexin, a heme scavenger protein, as observed in PM infants could also increase heme toxicity. This model of cell-free hemoglobin/heme toxicity in the lung is supported by known adverse effects of intravascular hemolysis in liver, spleen, kidney and vessels that are associated with sickle cell disease, malaria, blood transfusions cardiopulmonary bypass and sepsis [31].

The reason for the observed decrease in TA hemoglobin, but not in other candidate biomarker proteins, for most infants in our time course studies is unknown. Because albumin and total IgG did not decrease with time, the decline in hemoglobin levels in lung fluid cannot be explained by improving overall barrier function. In general, lung disease does not improve in these infants during this time and there were no consistent changes in respiratory support or medications with one exception: iNO was started in all infants after collection of the first TA sample. We speculate that NO therapy may reduce lung injury secondary to hemolysis similar to the observed benefit of iNO for acute kidney injury associated with bypass [34].

If a role of cell-free hemoglobin in the pathogenesis of PM in infants is confirmed by further studies, some possible targeted therapeutic approaches can be considered. In addition to iNO treatment, which reduces BPD in infants of maternally-identified African-American women [35], therapies with scavenger proteins such as haptoglobin and hemopexin could reduce the concentration of hemoglobin and heme, respectively [31]. One approach to address oxidative stress is superoxide dismutase, a therapy that has been tested in infants [36].

Another candidate protein of particular interest with regard to pathogenesis of BPD is vitamin D binding protein (GC), which was lower in TA of PM infants. Vitamin D binding protein is the major transporter for vitamin D, a scavenger of actin in damaged cells as occur in the injured lung [37], implicated as a positive regulator of lung development [38], and impacts the immune response by enhancing complement-mediated neutrophil chemotaxis. Altered vitamin D binding protein has been associated with pulmonary sarcoidosis, acute respiratory distress syndrome, chronic obstructive pulmonary disease and asthma in adults [39, 40], and there is one report for association with BPD [41].

In summary, we provide new information on the protein composition of lung fluid in premature infants with lung disease, combining global proteomic and transcriptomic analyses and using a cohort of infants with extensive clinical data. The findings allow assignment of the origin of most lung fluid proteins as plasma or lung cells and also identify candidate biomarker proteins for continuing infant lung disease through the first year of life. We observed elevated hemoglobin subunits in PM infants and discuss a possible causal role for these proteins in the pathogenesis of lung disease.

## Supporting information

**S1 Fig. Assay variability based on peptide count.** A, Intra-assay. B, Inter-assay. The coefficient of variability (CV) from replicate experiments improves with increasing peptide count. A peptide count >3 was used for all analyses.  
(TIF)

**S1 Table. Proteins detected in infant tracheal aspirate, sorted by AI.**  
(XLSX)

**S2 Table. Abundance of fetal lung mRNA for TA proteins, sorted by mean RNA cpm.** Data by RNAseq for human fetal lung. ND, not detected, mean cpm <4.  
(XLSX)

**S3 Table. Cellular localization of TA proteins by gene ontology.** GO categories for TA proteins with detectable mRNA in fetal lung. p values <0.05 with Bonferonni correction.  
(XLSX)

**S4 Table. Biological process of TA proteins by gene ontology.** GO categories for TA proteins with detectable mRNA in fetal lung. p values <0.05 with Bonferonni correction.  
(XLSX)

**S5 Table. Proteins detected in medium of human fetal lung explants, sorted by AI.** Data for culture medium from explants of 3 specimens of human fetal lung.

(XLSX)

**S6 Table. Proteins detected in infant plasma, sorted by peptide count.** Data from plasma of 8 premature infants.

(XLSX)

**S7 Table. TA protein abundance for PM vs No PM infants, sorted by fold difference.**

(XLSX)

## Acknowledgments

We thank Shreya Chand for technical assistance and all the TOLSURF Investigators, study coordinators, physicians, nurses, respiratory therapists and the families who participated in the TOLSURF study.

## Author Contributions

**Conceptualization:** Philip L. Ballard, Alma L. Burlingame.

**Data curation:** Juan Oses-Prieto, Cheryl Chapin.

**Formal analysis:** Mark R. Segal.

**Funding acquisition:** Philip L. Ballard, Alma L. Burlingame.

**Investigation:** Juan Oses-Prieto, Cheryl Chapin.

**Project administration:** Philip L. Ballard.

**Resources:** Philip L. Ballard, Roberta A. Ballard.

**Software:** Juan Oses-Prieto.

**Supervision:** Philip L. Ballard, Alma L. Burlingame.

**Visualization:** Cheryl Chapin.

**Writing – original draft:** Philip L. Ballard.

**Writing – review & editing:** Juan Oses-Prieto, Cheryl Chapin, Mark R. Segal, Roberta A. Ballard, Alma L. Burlingame.

## References

1. Jobe AH, Bancalari E. Bronchopulmonary dysplasia. *Am J Respir Crit Care Med.* 2001; 163(7):1723–9. <https://doi.org/10.1164/ajrccm.163.7.2011060> PMID: 11401896
2. Gibson AM, Reddington C, McBride L, Callanan C, Robertson C, Doyle LW. Lung function in adult survivors of very low birth weight, with and without bronchopulmonary dysplasia. *Pediatr Pulmonol.* 2015; 50(10):987–94. <https://doi.org/10.1002/ppul.23093> PMID: 25195792
3. Greenough A, Ahmed N. Perinatal prevention of bronchopulmonary dysplasia. *Journal of perinatal medicine.* 2013; 41(1):119–26. <https://doi.org/10.1515/jpm-2012-0084> PMID: 23096098
4. Bhandari A, Bhandari V. Biomarkers in bronchopulmonary dysplasia. *Paediatric respiratory reviews.* 2013; 14(3):173–9. <https://doi.org/10.1016/j.prrv.2013.02.008> PMID: 23523392
5. Ballard PL, Keller RL, Truog WE, Chapin C, Horneman H, Segal MR, et al. Surfactant status and respiratory outcome in premature infants receiving late surfactant treatment. *Pediatr Res.* 2019; 85:305–11. <https://doi.org/10.1038/s41390-018-0144-3> PMID: 30140069
6. Yang M, Kohler M, Heyder T, Forsslund H, Garberg HK, Karimi R, et al. Long-term smoking alters abundance of over half of the proteome in bronchoalveolar lavage cell in smokers with normal spirometry,

- with effects on molecular pathways associated with COPD. *Respir Res.* 2018; 19(1):40. <https://doi.org/10.1186/s12931-017-0695-6> PMID: 29514648
7. Martinez-Bravo MJ, Wahlund CJ, Qazi KR, Moulder R, Lukic A, Radmark O, et al. Pulmonary sarcoidosis is associated with exosomal vitamin D-binding protein and inflammatory molecules. *J Allergy Clin Immunol.* 2017; 139(4):1186–94. <https://doi.org/10.1016/j.jaci.2016.05.051> PMID: 27566455
  8. Bhargava M, Viken KJ, Dey S, Steinbach MS, Wu B, Jagtap PD, et al. Proteome Profiling in Lung Injury after Hematopoietic Stem Cell Transplantation. *Biol Blood Marrow Transplant.* 2016; 22(8):1383–90. <https://doi.org/10.1016/j.bbmt.2016.04.021> PMID: 27155584
  9. Foster MW, Morrison LD, Todd JL, Snyder LD, Thompson JW, Soderblom EJ, et al. Quantitative proteomics of bronchoalveolar lavage fluid in idiopathic pulmonary fibrosis. *Journal of proteome research.* 2015; 14(2):1238–49. <https://doi.org/10.1021/pr501149m> PMID: 25541672
  10. Bhargava M, Becker TL, Viken KJ, Jagtap PD, Dey S, Steinbach MS, et al. Proteomic profiles in acute respiratory distress syndrome differentiates survivors from non-survivors. *PLoS One.* 2014; 9(10): e109713. <https://doi.org/10.1371/journal.pone.0109713> PMID: 25290099
  11. Nguyen EV, Gharib SA, Palazzo SJ, Chow YH, Goodlett DR, Schnapp LM. Proteomic profiling of bronchoalveolar lavage fluid in critically ill patients with ventilator-associated pneumonia. *PLoS One.* 2013; 8(3):e58782. <https://doi.org/10.1371/journal.pone.0058782> PMID: 23505561
  12. Nguyen EV, Gharib SA, Crothers K, Chow YH, Park DR, Goodlett DR, et al. Proteomic landscape of bronchoalveolar lavage fluid in human immunodeficiency virus infection. *Am J Physiol Lung Cell Mol Physiol.* 2014; 306(1):L35–42. <https://doi.org/10.1152/ajplung.00140.2013> PMID: 24213920
  13. Magagnotti C, Matassa PG, Bachi A, Vendettuoli V, Fermo I, Colnaghi MR, et al. Calcium signaling-related proteins are associated with broncho-pulmonary dysplasia progression. *Journal of proteomics.* 2013; 94:401–12. <https://doi.org/10.1016/j.jprot.2013.10.007> PMID: 24140977
  14. Ballard RA, Keller RL, Black DM, Ballard PL, Merrill JD, Eichenwald EC, et al. Randomized Trial of Late Surfactant Treatment in Ventilated Preterm Infants Receiving Inhaled Nitric Oxide. *J Pediatr.* 2016; 168:23–9 e4. <https://doi.org/10.1016/j.jpeds.2015.09.031> PMID: 26500107
  15. Keller RL, Eichenwald EC, Hibbs AM, Rogers EE, Wai KC, Black DM, et al. The Randomized, Controlled Trial of Late Surfactant: Effects on Respiratory Outcomes at 1-Year Corrected Age. *J Pediatr.* 2017; 183:19–25 e2. <https://doi.org/10.1016/j.jpeds.2016.12.059> PMID: 28100402
  16. Ballard RA, Truog WE, Cnaan A, Martin RJ, Ballard PL, Merrill JD, et al. Inhaled nitric oxide in preterm infants undergoing mechanical ventilation. *N Engl J Med.* 2006; 355(4):343–53. <https://doi.org/10.1056/NEJMoa061088> PMID: 16870913
  17. Barrette AM, Roberts JK, Chapin C, Egan EA, Segal MR, Oses-Prieto JA, et al. Antiinflammatory Effects of Budesonide in Human Fetal Lung. *Am J Respir Cell Mol Biol.* 2016; 55(5):623–32. <https://doi.org/10.1165/rcmb.2016-0068OC> PMID: 27281349
  18. Gonzales LW, Guttentag SH, Wade KC, Postle AD, Ballard PL. Differentiation of human pulmonary type II cells in vitro by glucocorticoid plus cAMP. *Am J Physiol Lung Cell Mol Physiol.* 2002; 283(5): L940–51. <https://doi.org/10.1152/ajplung.00127.2002> PMID: 12376347
  19. Guan S, Price JC, Prusiner SB, Ghaemmaghami S, Burlingame AL. A data processing pipeline for mammalian proteome dynamics studies using stable isotope metabolic labeling. *Molecular & cellular proteomics: MCP.* 2011; 10(12):M111.010728. <https://doi.org/10.1074/mcp.M111.010728> PMID: 21937731
  20. Clauser KR, Baker P, Burlingame AL. Role of accurate mass measurement (+/- 10 ppm) in protein identification strategies employing MS or MS/MS and database searching. *Analytical chemistry.* 1999; 71(14):2871–82. <https://doi.org/10.1021/ac9810516> PMID: 10424174
  21. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* 2019; 47(D1):D442–D450. <https://doi.org/10.1093/nar/gky1106> PMID: 30395289
  22. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature genetics.* 2000; 25(1):25–9. PMID: 10802651
  23. Wade KC, Guttentag SH, Gonzales LW, Maschhoff KL, Gonzales J, Kolla V, et al. Gene Induction during Differentiation of Human Pulmonary Type II Cells In Vitro. *Am J Respir Cell Mol Biol.* 2006; 34:727–37. <https://doi.org/10.1165/rcmb.2004-0389OC> PMID: 16474099
  24. Suski M, Bokinić R, Szwarc-Duma M, Madej J, Bujak-Gizycka B, Borszewska-Kornacka MK, et al. Plasma proteome changes in cord blood samples from preterm infants. *J Perinatol.* 2018; 38(9):1182–9. <https://doi.org/10.1038/s41372-018-0150-7> PMID: 29910464
  25. Holter JF, Weiland JE, Pacht ER, Gadek JE, Davis WB. Protein permeability in the adult respiratory distress syndrome. Loss of size selectivity of the alveolar epithelium. *J Clin Invest.* 1986; 78(6):1513–22. <https://doi.org/10.1172/JCI112743> PMID: 2431002



26. Pereira-Fantini PM, Pang B, Byars SG, Oakley RB, Perkins EJ, Dargaville PA, et al. Preterm Lung Exhibits Distinct Spatiotemporal Proteome Expression at Initiation of Lung Injury. *Am J Respir Cell Mol Biol*. 2019; 61(5):631–42. <https://doi.org/10.1165/rcmb.2019-0084OC> PMID: 30995072
27. Pereira-Fantini PM, Byars SG, McCall KE, Perkins EJ, Oakley RB, Dellaca RL, et al. Plasma proteomics reveals gestational age-specific responses to mechanical ventilation and identifies the mechanistic pathways that initiate preterm lung injury. *Sci Rep*. 2018; 8(1):12616. <https://doi.org/10.1038/s41598-018-30868-x> PMID: 30135517
28. Ballard PL, Lee JW, Fang X, Chapin C, Allen L, Segal MR, et al. Regulated gene expression in cultured type II cells of adult human lung. *Am J Physiol Lung Cell Mol Physiol*. 2010; 299(1):L36–50. <https://doi.org/10.1152/ajplung.00427.2009> PMID: 20382749
29. Olmeda B, Umstead TM, Silveyra P, Pascual A, Lopez-Barneo J, Phelps DS, et al. Effect of hypoxia on lung gene expression and proteomic profile: insights into the pulmonary surfactant response. *Journal of proteomics*. 2014; 101:179–91. <https://doi.org/10.1016/j.jprot.2014.02.019> PMID: 24576641
30. Collard KJ. Transfusion related morbidity in premature babies: Possible mechanisms and implications for practice. *World J Clin Pediatr*. 2014; 3(3):19–29. <https://doi.org/10.5409/wjcp.v3.i3.19> PMID: 25254181
31. Schaer DJ, Buehler PW, Alayash AI, Belcher JD, Vercellotti GM. Hemolysis and free hemoglobin revisited: exploring hemoglobin and heme scavengers as a novel class of therapeutic proteins. *Blood*. 2013; 121(8):1276–84. <https://doi.org/10.1182/blood-2012-11-451229> PMID: 23264591
32. Steinhorn RH. Therapeutic approaches using nitric oxide in infants and children. *Free radical biology & medicine*. 2011; 51(5):1027–34. <https://doi.org/10.1016/j.freeradbiomed.2011.01.006> PMID: 21237265
33. Shaver CM, Upchurch CP, Janz DR, Grove BS, Putz ND, Wickersham NE, et al. Cell-free hemoglobin: a novel mediator of acute lung injury. *Am J Physiol Lung Cell Mol Physiol*. 2016; 310(6):L532–41. <https://doi.org/10.1152/ajplung.00155.2015> PMID: 26773065
34. Lei C, Berra L, Rezoagli E, Yu B, Dong H. Nitric oxide decreases acute kidney injury and stage 3 chronic kidney disease after cardiac surgery. *Am J Respir Crit Care Med*. 2018; 198:1279–87. <https://doi.org/10.1164/rccm.201710-2150OC> PMID: 29932345
35. Askie LM, Davies LC, Schreiber MD, Hibbs SM, Ballard PL, Ballard RA. Race Effects of Inhaled Nitric Oxide in Preterm Infants: An Individual Participant Data Meta-Analysis. *J Pediatr*. 2018; 193:34–9. <https://doi.org/10.1016/j.jpeds.2017.10.004> PMID: 29241680
36. Davis JM, Parad RB, Michele T, Allred E, Price A, Rosenfeld W. Pulmonary outcome at 1 year corrected age in premature infants treated at birth with recombinant human CuZn superoxide dismutase. *Pediatrics*. 2003; 111(3):469–76. <https://doi.org/10.1542/peds.111.3.469> PMID: 12612223
37. Chishimba L, Thickett DR, Stockley RA, Wood AM. The vitamin D axis in the lung: a key role for vitamin D-binding protein. *Thorax*. 2010; 65(5):456–62. <https://doi.org/10.1136/thx.2009.128793> PMID: 20435872
38. Lykkedegn S, Sorensen GL, Beck-Nielsen SS, Christesen HT. The impact of vitamin D on fetal and neonatal lung maturation. A systematic review. *Am J Physiol Lung Cell Mol Physiol*. 2015; 308(7):L587–602. <https://doi.org/10.1152/ajplung.00117.2014> PMID: 25595644
39. Powe CE, Evans MK, Wenger J, Zonderman AB, Berg AH, Nalls M, et al. Vitamin D-binding protein and vitamin D status of black Americans and white Americans. *N Engl J Med*. 2013; 369(21):1991–2000. <https://doi.org/10.1056/NEJMoa1306357> PMID: 24256378
40. Malik S, Fu L, Juras DJ, Karmali M, Wong BY, Gozdzik A, et al. Common variants of the vitamin D binding protein gene and adverse health outcomes. *Crit Rev Clin Lab Sci*. 2013; 50(1):1–22. <https://doi.org/10.3109/10408363.2012.750262> PMID: 23427793
41. Serce Pehlevan O, Karatekin G, Koksall V, Benzer D, Gursoy T, Yavuz T, et al. Association of vitamin D binding protein polymorphisms with bronchopulmonary dysplasia: a case-control study of gc globulin and bronchopulmonary dysplasia. *J Perinatol*. 2015; 35(9):763–7. <https://doi.org/10.1038/jp.2015.58> PMID: 26067474